Detecting Temporal Trends in Species Assemblages with Bootstrapping Procedures and Hierarchical Models

Detecting Temporal Trends In Species Assemblages With Bootstrapping Procedures and Hierarchical Models

Nicholas J. Gotelli[1,*]

Robert M. Dorazio[2]

Aaron M. Ellison[3]

Gary D. Grossman[4]

[1]Department of Biology
University of Vermont
Burlington, VT 05405 USA

[2]Southeast Ecological Science Center
U.S. Geological Survey
& Department of Statistics
University of Florida
Gainesville, FL 32611 USA

[3] Harvard Forest
Harvard University
Petersham, MA 01366 USA

[4]Warnell School of Forestry & Natural Resources
University of Georgia
Athens, GA, USA 30602

[*]Correspondence author: ngotelli@uvm.edu

ABSTRACT

Quantifying patterns of temporal trends in species assemblages is an important analytical challenge in community ecology. We describe methods of analysis that can be applied to a matrix of counts of individuals that is organized by species (rows) and time-ordered sampling periods (columns). We first developed a bootstrapping procedure to test the null hypothesis of random sampling from a stationary species abundance distribution with temporally varying sampling probabilities. This procedure can be modified to account for undetected species. We next developed a hierarchical model to estimate species-specific trends in abundance while accounting for species-specific probabilities of detection. We analyzed two long-term data sets on stream fishes and grassland insects to demonstrate these methods. For both assemblages, the bootstrap test indicated that temporal trends in abundance were more heterogeneous than expected under the null model. We used the hierarchical model to estimate trends in abundance and identified sets of species in each assemblage that were steadily increasing, decreasing, or remaining constant in abundance over more than a decade of standardized annual surveys. Our methods of analysis are broadly applicable to other ecological data sets, and they represent an advance over most existing procedures, which do not incorporate effects of incomplete sampling and imperfect detection.

KEY WORDS: Temporal trends, species abundance, null model, hierarchical model, stream fishes, grassland insects

62 INTRODUCTION

63 Quantifying change in the structure of plant and animal communities is an important challenge for
64 ecology in the 21st century (Walther *et al.* 2002). Species composition and abundance can respond
65 directly to long-term changes in abiotic factors (Dunson & Travis 1991) and indirectly to changes in the
66 occurrence or abundance of other species (White *et al.* 2006). Dramatic and rapid changes in
67 community structure may result from the addition or loss of keystone species (Mills *et al.* 1993),
68 foundation species (Ellison *et al.* 2005), ecosystem engineers (Jones *et al.* 1994), and some (but not all)
69 non-native species (Manchester & Bullock 2000). Other changes may be more subtle, because the
70 abundance of individual species can gradually increase or decrease over long periods of time, as in
71 scenarios of a "shifting baseline" (Pauly 1995). Long-term trends may be difficult to detect because of
72 substantial short-term noise and variability in abundances between consecutive samples.

73 However, not all observed changes in community structure through time are biologically relevant. Most
74 measures of community structure and diversity are sensitive to sampling effort and to the number of
75 individuals counted (Gotelli & Colwell 2001). These quantities are rarely constant through time, even
76 with standardized monitoring programs. Rare species, in particular, are expected to occur more often
77 when sampling is more thorough (Chao *et al.* 2009). Even the appearance of a previously unrecorded
78 species need not signal a true change in community structure. Biodiversity sampling is labor intensive
79 and notoriously incomplete (Lawton *et al.* 1998), and "new" species occurrence records— especially of
80 plants and invertebrates— are routinely made, even in sites in that have been well-sampled for many
81 years (e.g. Longino *et al.* 2002).

82 A variety of univariate and multivariate methods have been proposed to quantify the degree of
83 community change through time (Collins *et al.* 2000, Fujiwara & Mohr 2009), and to detect temporal
84 trends in community structure (Clarke 1993, Solow 1994). However, with few exceptions (*e.g.*, Dorazio
85 *et al.* 2010), existing methods do not account for incomplete sampling and imperfect detection. Instead,

most methods assume that the absence of a species from a sampling period represents a "true" zero, and not a detection error (Royle & Dorazio 2008). Most procedures also ignore species that may have been present in a region, but were never detected in any of the samples (Colwell & Coddington 1994).

In this study, we develop new methods for quantifying temporal trends in species abundances that account for errors in detection of individuals. Our methods are appropriate for analyzing species-specific counts of individuals recorded from repeated surveys of a single site. We first develop a bootstrap procedure for testing a null hypothesis in which the counts are assumed to have arisen from sampling a stationary distribution of relative species abundances with temporally varying sampling probabilities. We then develop a hierarchical model of the counts to estimate species-specific trends in abundance while accounting for species-specific probabilities of detection. Both methods of analysis are illustrated for two long-term data sets on stream fishes and grassland insects.

MATERIALS & METHODS

DATA STRUCTURE

The data for our analyses may be organized as a matrix of counts of individuals of S species (rows) recorded during T successive sampling periods (columns). The matrix entry y_{ij} is the number of individuals of species i that were observed at sampling time j ($i = 1,\ldots, S$; $j = 1,\ldots, T$). This simple data structure arises in many ecological studies in which species assemblages are repeatedly sampled at a site. Although the samples do not have to be evenly spaced in time, our methods are intended for analysis of long-term trends in abundance, not for short-term or periodic changes in abundance (e.g., seasonality). We illustrate our methods with two data sets: a 13-year record of annual counts of 55 fish species seined from a mid-western U.S.A. stream (Grossman *et al.* 1982), and a 14-year record of annual counts of 9 insect species collected from sticky traps in a successional grassland plot at the U.S.A. Kellogg Biological Station (KBS 1995). Table 1 summarizes sampling details for each of these

studies.

The sampling design and collecting methods for the stream fish study have been described previously (Whitaker 1976; Grossman *et al.* 1982, 1985) and are only summarized here. A 120 m long × 23 m wide section of Otter Creek, Vigo County, Indiana, USA, was surveyed annually between 1962 and 1974. The site contained a diversity of substrata and depths and can be considered representative of many streams in the midwestern United States. During the study period, the site retained a relatively stable physical structure. Fishes were sampled using a seine, and all collections were supervised by a single investigator, so effort was relatively consistent. There was some minor variation present in sampling efficiency produced by differences in stream depth among years, although the investigators always attempted to keep the area sampled constant. All fishes captured were identified to species and counted, except in a few cases in which a species was extremely abundant. In those cases, numerical estimates were derived from subsamples of the total catch. All fishes were immediately returned live to the site, except for voucher specimens, which were preserved for later identification (Grossman *et al.* 1982, 1985).

The insect data were collected as part of the Long Term Ecological Research (LTER) network (LTER 2007) at the Kellogg Biological Station (KBS) in northern Michigan, U.S.A. At KBS, a set of 7 1 ha crop rotational treatments have been replicated in 6 blocks on a single 60 ha plot (KBS 1995). We used data from Treatment 7, a native successional treatment that was abandoned after spring plowing in 1989. The plots are surveyed with yellow sticky traps that are replaced weekly from May to August. These traps collect insect predators, many of which are identified only to family (Chrysopidae and Lampyridae). We used data for 9 taxa that were identified to species and were sampled in all years of the study. Within each sampling season, we pooled data from all sticky traps and all plots to generate annual counts for each species.

NULL MODEL ANALYSIS

In this section we describe a null model for detecting temporal trends in species' abundances. We use the term "null model" to represent a model wherein an S x T matrix of counts of individuals is assumed to arise by randomly selecting individuals from the species assemblage according to each species' relative abundance and a set of temporally varying sampling probabilities. The key issue is that no particular ecological process or mechanism is assumed to have generated the matrix of counts; thus the model incorporates simple sampling effects, but is "null" with respect to processes that might induce trends in species abundance (Gotelli & Graves 1996). The total number of individuals of species i in all sampling periods is:

$$m_i = \sum_{j=1}^{T} y_{ij} \tag{1}$$

The total number of individuals of all species observed during sampling period j is:

$$n_j = \sum_{i=1}^{S} y_{ij} \tag{2}$$

Let N equal the total number of individuals summed across all species and samples:

$$N = \sum_{i=1}^{S}\sum_{j=1}^{T} y_{ij} = \sum_{i=1}^{S} m_i = \sum_{j=1}^{T} n_j \tag{3}$$

We define the relative abundance of species i in the source pool of N individuals as:

$$s_i = \frac{m_i}{N} \tag{4}$$

Similarly, we define the relative sampling intensity during the jth survey as:

$$q_j = \frac{n_j}{N} \tag{5}$$

Under the assumptions of the null model, s_i is regarded as the probability that an individual drawn from the source pool of N individuals belongs to species i, and q_j is regarded as the probability that an individual is observed in the jth sampling period, regardless of species.

To conduct a bootstrap test of the null model, we first randomly assign each of the N individuals in the total collection to a particular sample, with probability q_j. Once all the individuals are assigned, we then assign them species identities by sampling randomly with replacement from the distribution of s_i values. This two-step process does not depend on the order of conditioning; the same distribution would be obtained by first assigning individuals to species using the s_i values, and then assigning these individuals to particular samples using the q_j values.

This null model describes a multinomial sampling process that is conditional on N, the total number of individuals observed. The simulated number of individuals y_{ij} of species i in sample j depends on s_i, the proportional representation of species i in the source pool, q_j, the proportion of individuals sampled at time j, and N. The null hypothesis is that variability among species in temporal trends is no greater than would be expected from this simple model of sampling with replacement from an underlying stationary distribution of relative abundance. The alternative hypothesis is that at least some species in the assemblage are systematically increasing or decreasing, leading to changes in relative abundance that cannot be accounted for entirely by sampling effects.

The null hypothesis is defined by the following hierarchical model:

$$n_1, n_2, \ldots, n_T \sim \text{multinomial}(N, q_1, q_2, \ldots, q_T) \quad (6)$$

$$y_{1j}, y_{2j}, \ldots, y_{Sj} \sim \text{multinomial}(n_j, p_1, p_2, \ldots, p_S) \quad (7)$$

and our randomization procedure is entirely consistent with this model. Based on this model, the marginal distribution of the counts is multinomial:

$$y_{11}, y_{21}, \ldots, y_{ST} \sim \text{multinomial}(N, p_1 q_1, p_2 q_1, \ldots, p_S q_T) \quad (8)$$

Therefore, under the null hypothesis the expected value of each count is proportional to the product of species relative abundance and year-specific sampling probability.

To estimate temporal trends from the observed data, we first fit a simple linear model to the count data for each species *i*:

$$y_{ij} = \beta_{0i} + \beta_{1i} t_j + \varepsilon_{ij} \quad (9)$$

where t_j is time (in arbitrary units of years, months, or time-steps), β_{0i} is the intercept, β_{1i} is the slope of the regression for species *i*, and the error term ε_{ij} has a normal distribution $\left(\varepsilon_{ij} \sim N(0, \sigma_i^2)\right)$.

We are interested in β_{1i}, because it measures the simple temporal trend in abundance for species *i*. Temporal change (*TC*) in the entire assemblage can be then be quantified as the sample variance of the estimated β_{1i} values:

$$TC = \frac{1}{S-1} \sum_{i=1}^{S} \left(\beta_{1i} - \bar{\beta}_1\right)^2 \quad (10)$$

The larger *TC*, the more heterogeneity there is in the temporal trends of the component species, and the more change in composition of the assemblage that will be seen at future sampling dates. As described below, the number of species generated in the null assemblages was not constant. However, for both the real and the simulated matrices, *TC* was calculated only for species that were present at least once in the matrix. Following standard procedures for resampling tests (Manly 2009), we generated 1000 null assemblages, and calculated *TC* for each. We estimate the probability of obtaining *TC* if the null

198 hypothesis were true by comparing the observed TC to the histogram of simulated TC values.

199 Because the results are potentially sensitive to the assumption of simple linear trends in y_{ij}, we fit two
200 alternative regression models based on log-log and log-linear transformations of (y_{ij} + 1) and t_j. The
201 same transformations were applied to the real and the simulated data. Although these alternative models
202 incorporated non-linear trends in species temporal trajectories, the transformations had no qualitative
203 effect on the outcome of the null model tests. Therefore, we present results only from analyses of the
204 untransformed data fit with a linear trend line.

205 UNDETECTED SPECIES

206 The construction of the null matrix is similar to a simulation of rarefaction (Sanders 1969, Hurlbert
207 1971), in which a small assemblage is simulated by random draws of subsamples of n_j individuals from
208 the larger sample of N. However, in rarefaction, sampling is done without replacement (Simberloff
209 1978). Because our null model treats the source pool as a permanent stationary distribution, we sampled
210 from it with replacement. In practice, the results will not differ unless the sample sizes are so small that
211 n_j is a relatively large fraction of N, which is not the case for these data sets. Rarefaction also conditions
212 on n_j, the observed count in a particular sample, whereas our multinomial model conditions on N, the
213 total number of individuals.

214 This procedure implicitly addresses detection error because species (especially rare ones) that are
215 present in the aggregated collection N may not be represented in any particular sample n_j. In some null
216 assemblages, species that were very rare in the original data set may be missing from all n_j samples.
217 Because biodiversity sampling is notoriously incomplete, there are also likely to be rare species in the
218 assemblage that were never encountered in the original samples (Colwell and Coddington 1994). We
219 expanded our null model to incorporate these undetected species. We first estimated the minimum
220 number of undetected species, $\hat{S}_1$, using a bias-corrected version of the familiar Chao2 estimator (Chao

1984; Equation (4) in Colwell 2009):

$$\hat{S} = \left(\frac{T-1}{T}\right)\left(\frac{Q_1(Q_1-1)}{2(Q_2+1)}\right) \quad (11)$$

where Q_1 is the number of species represented in exactly 1 time period ("uniques"), Q_2 is the number of species represented in exactly two time periods ("duplicates"), and T is the number of samples. The Chao2 index estimates the number of undetected species in the entire assemblage, not the number that may be undetected in any single sample. For the stream fish matrix, the estimated number of undetected species (rounded to the nearest whole integer) was 16. For the insect matrix, sampling was restricted to nine common species, and the estimated number of undetected species was 0.

Once the number of undetected species was estimated, it was necessary to assign them each a relative abundance s_i, so they could be included in the simulation. Estimating these s_i values would require knowledge of the precise form of the species abundance distribution, a long-standing unsolved problem in ecology (McGill *et al.*2007). As a simplifying first approximation, we assumed that s_i for each undetected species was equal to 0.5* s_i for the least abundant species observed in the assemblage. The reasoning is that if any of these undetected species occurred at a frequency greater than this, they would likely have been detected at least once in the original sample. For the stream fish data, s_i for each of the 16 undetected species was set at 3.414135×10^{-5}. Because many of the undetected species are probably much more rare than this, our procedure allows for the greatest possible influence of undetected species. Nevertheless, the results for the stream fish matrix were identical with and without the inclusion of undetected species. However, because the observed number of species is always a biased under-estimator of true species richness, we present the full analyses here with the undetected species included in the null model.

If the observed value of TC is larger than those of 950 of the 1000 simulated TC values ($p < 0.05$, one-

tailed test), then the temporal trends in the set of observed species are more heterogeneous than can be accounted for by the null model: at least some species are either increasing or decreasing more rapidly than would be expected from sampling effects and undetected species. The null model was programmed and implemented in the statistical language R (R Core Development Team, 2008; Appendix A).

Hierarchical Model of Trend in Abundances

The null model provides a simple test for heterogeneity in species trends. If this null hypothesis is rejected, the next step is to estimate the rate of change in abundance of each species. As before, y_{ij} is the count of species i in sample j. We assume that the count y_{ij} depends on the abundance N_{ij} present during the jth survey and on each individual's probability of capture p_{ij}, as follows:

$$y_{ij} | N_{ij}, p_{ij} \sim \text{binomial}(N_{ij}, p_{ij}) \quad (12)$$

To estimate trend, we assume population abundances can be described as:

$$N_{ij} | \lambda_{ij} \sim \text{Poisson}(\lambda_{ij}) \quad (13)$$

$$\lambda_{ij} = \lambda_{i0} \exp(r_i t_j) \quad (14)$$

where λ_{ij} denotes mean abundance of species i during survey j and where t_j denotes the year of the jth survey. Trend in λ_{ij} values is specified using the familiar exponential growth model (Eq. 14), which includes a species-specific intercept parameter λ_{i0} and a net population growth rate parameter r_i.

Note that N_{ij} is not actually observed. N_{ij} is a parameter of the model that represents the number of individuals of species i which are present and available to be captured during the jth survey. The observation y_{ij} can be interpreted as a negatively biased estimator of N_{ij}, with the level of bias depending on the magnitude of p_{ij}, the unknown probability of capture for individuals of species i.

In the absence of replicated observations, we cannot estimate temporal changes in both N_{ij} and p_{ij}.

Therefore, we assume that capture probabilities vary among species but not among surveys (i.e., we assume $p_{ij} = p_i$). Even with this simplifying assumption, the hierarchical model composed of Eqs. 12–14 contains more parameters than can be estimated from the data. To solve this problem, N_{ij} may be eliminated from the model by integrating the joint distribution of y_{ij} and N_{ij}. This integration can be done analytically to obtain the following marginalized version of the hierarchical model:

$$y_{ij} | p_i, \lambda_{i0}, r_i \sim \text{Poisson}\left(p_i \lambda_{i0} \exp(r_i t_j)\right) \quad (15)$$

Note that this model may be viewed conceptually as a Poisson regression model. For example, let μ_{ij} denote the Poisson mean for y_{ij}. The logarithm of μ_{ij} is a linear combination of the marginal model's parameters:

$$\log(\mu_{ij}) = \log(p_i) + \log(\lambda_{i0}) + r_i t_j \quad (16)$$

However, p_i and λ_{i0} are not identifiable parameters in Eq. 16 (i.e., both parameters cannot be estimated); therefore, we combine these parameters into a common regression intercept parameter (say, $a_i = \log(p_i \lambda_{i0})$) to obtain

$$\log(\mu_{ij}) = a_i + r_i t_j \quad (17)$$

From this equation, the T observations, $y_{i1}, y_{i2}, \ldots, y_{iT}$, can be used to estimate the parameters a_i and r_i. We are interested primarily in the latter parameter r_i, which denotes the trend in abundance of species i; however, our formulation of the intercept parameter a_i reveals explicitly the combined roles of mean abundance and capture probability in the model.

The model specified by Eqs. 15 and 17 can be fitted to each species separately. However, doing so may produce estimates of trend that are unstable or highly imprecise for species whose abundance appears to be low (as indicated by counts that contain several zeros and ones). Therefore, we extend the model as

follows:

$$r_i|\beta,\sigma \sim \text{normal}(\beta,\sigma^2) \quad (18)$$

where β denotes the average trend in abundance among species in this assemblage and σ denotes the level of variation in r_i values among species. This distributional assumption allows the counts of all species to be analyzed jointly so that information associated with species of moderate or high abundance can be used to stabilize the estimates of trend for species of low apparent abundance. Nevertheless, even with this assumption, there were not enough data to reliably estimate trends for very rare species that were represented by less than 10 individuals in the entire survey (25 of 55 stream fish species, and 2 of 9 insect species).

Equation 18 implies an exchangeability of trend parameters among species. This exchangeability formalizes the notion that although abundances may be increasing, decreasing, or constant for any particular species, each is also a member of a common assemblage. We expect that the temporal trends of the species in the stream fish assemblage are more similar to one another than they are to, say, the temporal trends of the species in the grassland insect assemblage. A restricted version of this model that corresponds to the null model assumes an identical growth rate $r_i = \beta$ for all species, so that $\sigma = 0$. We can fit this restricted model and compare it to the unrestricted model to assess whether the data support the null hypothesis that all species abundances have an identical trend.

METHOD OF ESTIMATION

The hierarchical model described by Eqs. 15, 16, and 18 may be fitted by maximizing the likelihood function obtained by integrating away the latent trend parameters. In our situation, however, this approach is counter-productive. In addition to the minor technical challenges of computing the integrals numerically, the trend parameters r_i are the quantities of primary scientific interest. Estimates of these

parameters and their uncertainties are actually needed to solve the inference problem. We therefore adopt a Bayesian approach to inference, which allows every parameter in the model to be estimated directly, including the species-specific trends in abundance.

In a Bayesian analysis, all inferences are based on the joint posterior distribution of model parameters. In our case the unnormalized, probability density function (pdf) of this distribution is

$$\pi(\boldsymbol{a}, \boldsymbol{r}, \beta, \sigma | \boldsymbol{Y}) \propto \pi(\beta, \sigma, \boldsymbol{a}) \prod_{i=1}^{S} g(r_i | \beta, \sigma) \prod_{j=1}^{T} f\left(y_{ij} | exp(a_i + r_i t_j)\right) \qquad (19)$$

where $\boldsymbol{a} = (a_1, \ldots, a_n)', \boldsymbol{r} = (r_1, \ldots, r_n)'$, and $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)'$. Here, g(•|β, σ) denotes the pdf of a normal distribution with mean β and variance σ^2 , f(•|μ_{ij}) denotes the probability mass function of a Poisson distribution with mean μ_{ij}, and π(β, σ, a) denotes the pdf of the prior distribution of the parameters β, σ, and a.

The posterior pdf cannot be written in closed form owing to the analytically intractable integrals in the normalizing constant (not shown in Eq. 19). Therefore, we estimated the model's parameters using Markov chain Monte Carlo algorithms (Robert and Casella, 2004) to obtain an arbitrarily large sample of the joint posterior distribution. Specifically, we fit the model using the WinBUGS software (Lunn *et al.* 2000), which is an implementation of the BUGS language for specifying models and doing Bayesian analyses (Gilks *et al.*, 1994).

To obtain the posterior sample, we assumed a set of mutually independent noninformative prior distributions for β, σ, and a. We assumed normal(0, 100^2) priors for β and a_i and a uniform(0, 10) prior for σ. Each of five Markov chains was independently initialized and computed for a total of 21,000 draws. The first 1000 draws in each chain were discarded as "burn-in", and every fifth draw in the rest

of each chain was retained to form the posterior sample. Consequently, these calculations yielded a posterior sample of 20,000 draws, which was used to compute estimates of the model's parameters and their 95% credible intervals (Appendix B).

RESULTS

NULL MODEL ANALYSIS

For the stream fish data, there was a non-significant decreasing trend in total abundance (Figure 1), caused primarily by extremely high abundances in the November 1966 sample ($n_8 = 5344$ individuals). For the null model analysis, this decreasing trend leads to the expectation of negative slopes for individual species, with a moderate amount of variation among species (Figure 2, left panel). However, the observed slopes were much more heterogeneous than this expectation: several species showed sharply increasing or decreasing trend lines (Figure 2, right panel), and the observed TC index was larger than that of all 1000 simulated assemblages (Table 1).

For the insect data, there was a marginally non-significant increasing trend in total abundance (Figure 3), with systematically greater abundances during the final sampling years. For the null assemblages created from this matrix, most species had increasing trend lines (Figure 4, left panel). However, the observed slopes were again much more heterogeneous than expected (Figure 4, right panel). As with the stream fish data, the observed heterogeneity among slopes (TC) was greater than that of any of the simulated assemblages (Table 1).

TRENDS IN ABUNDANCES

For the stream fish data, the hierarchical model identified 7 species with significant increases in abundance, 17 species with significant declines in abundance, and 6 species with no significant trend (Figure 5). A negative estimate of average trend, $\hat{\beta} = -0.152$ (95% credible interval: (-0.289, -0.024)), also indicates that species with declining abundances outnumbered those with increasing abundances.

There is little doubt that trends in population abundance differed substantially among species. The posterior distribution of σ (Figure 6, left panel) provides virtually no support for the hypothesis that σ = 0.

For the grassland insect data, the hierarchical model identified 2 species with significant increases in abundance, 3 species with significant declines in abundance, and 2 species with no significant trend (Figure 7). The estimate of average trend , $\hat{\beta}$= -0.026 (95% credible interval: (-0.352, 0.297)), reflects the nearly equal numbers of species with increasing and decreasing abundances. As with the stream fish data, the posterior distribution of σ (Figure 6, right panel) does not support the null hypothesis (σ = 0) of identical trend lines for these 7 species.

DISCUSSION

The null model and the hierarchical model provide complementary information on temporal trends, and they both point to strong temporal re-organization of stream fish and insect grassland assemblages over periods greater than a decade. Because the insects were sampled in a successional plot, it is no surprise that strong temporal trends would be detected, as vegetation structure and arthropod prey assemblages were systematically changing through time. In fact, the two most rapidly increasing species, *Harmonia axiridis* and *Hippodamia glacialis*, never appeared in any of the traps until 6 and 7 years respectively after the sampling began. This is exactly the pattern that would be expected in a classic facilitation model of succession (Connell & Slatyer 1977). On the other hand, the abundance of the most common species in the samples, *Coccinella septempunctata* ($\bar{x}$ = 200 individuals/year), did not change significantly during the 14-year successional sequence (Figure 7).

The pattern for the stream fishes is more complex. Although no obvious physical changes were

observed in the habitat during the 15-year sampling period, 17 species showed significant declines, whereas only 7 species increased in abundance. The causal mechanisms behind these patterns are unclear because both generalist and specialist species were found in all categories, as were representatives of most North American taxonomic groups. Perhaps the preponderance of declining populations suggests that some species successfully invaded the site early in the time series, but were not able to sustain populations through local reproduction and began to decline. It is likely that flow variation plays some role in these trends, perhaps facilitating establishment of species in benign periods and causing substantial mortality during periods of high water (Grossman *et al.* 1982; 1998). High flow events may cause substantial mortality in stream fishes, especially if they occur during the reproductive period and destroy an entire year-class (Grossman *et al.* 1982; 1998). However, there was no evidence during the sampling period of declining flows or increased numbers of extreme flow events that might be linked to the decreasing abundance of 17 of the 55 species (Grossman & Sabo 2010). The decreasing trends in abundance of many stream fish species (Figure 5) are consistent with a shifting baseline scenario, but the causes of these declines are still unknown.

The results of both the null model and the hierarchical model are potentially sensitive to the functional form that is used to describe temporal trends. For the null model analysis, the results for these data sets were the same when the trends were fit with linear, semi-logarithmic, or log-log transformations of the original data. The estimated heterogeneity among species in temporal trends does not seem to be sensitive to the fitting procedure, perhaps because deviations caused by extreme sample numbers (such as the high counts in the stream fish data set in 1966) are also incorporated into the pattern in the null assemblages. Both the null model and the hierarchical model assume that species are independent of one another. However, it is unclear how the violation of this assumption (from strong species interactions) would systematically affect the estimates of temporal trends in abundance.

Because the hierarchical model is being used for parameter estimates of change (rather than just a simple dichotomous null model test), it is potentially more sensitive to violation of its assumptions. As we noted, one important assumption in this model is that capture probabilities are constant through time. Although this assumption may not be true, it probably matches the perspective of most field biologists, who typically assume that long-term monotonic changes in species counts with standardized sampling methods primarily reflect changes in abundance, rather than changes in detection or capture probabilities.

If species-specific capture probabilities are not constant, the magnitude of the deviations between observed and expected counts may be inflated. As long as these deviations do not vary systematically with time, the point estimates of trend will not be affected, although the credible intervals may be too narrow. Alternatively, if the deviations between expected and observed counts vary systematically with time, say changing from positive to negative values, the trend estimates will be very sensitive to an incorrect assumption of constant capture probability. For the data sets we analyzed, there was no evidence of a systematic lack of fit (Figure 8).

In the hierarchical model, the assumption of constant sampling probabilities was necessary only because of the extremely simple and unreplicated structure of the data matrix. With replication, it may be possible to estimate parameters for temporal trends in both abundance and detection probabilities. For example, the KBS insect data actually consist of weekly sticky trap counts collected from 6 replicated plots. Rather than pooling the data as we have done in this analysis, the individual trap records could be fit to a more complex hierarchical model (Royle & Dorazio 2008, Kery *et al.* 2009). The hierarchical model could also be expanded to incorporate species-specific covariates Z (such as body size, geographic range size, or degree of habitat specialization) that might be of interest for conservation purposes. Species-specific covariates could be used to model either the mean structure of

448 the elements of $\mathbf{r}$ in Eq. 18 or their covariances.

449

450 Both the bootstrap test and the hierarchical model assume that changes in abundance through time are
451 monotonic. If species show more complex patterns of temporal change (such as periodic fluctuations),
452 these could be accommodated by fitting polynomial or sine functions to the time series. However, at
453 least for these data sets, diagnostic analysis of residuals indicated little evidence for departures from
454 linearity over the time periods that were sampled. Moreover, a monotonic function is appropriate for
455 very short data series such as these ($T = 15$ samples for stream fishes and $T = 14$ samples for grassland
456 insects)

457

458 Finally, the frequent occurrence of rare species in natural assemblages continues to pose statistical
459 challenges. In the null model, all species, no matter how rare, are included in the analysis, and the null
460 assemblages even incorporate the possibility of species that were never detected in any of the samples.
461 In theory, rare species can contribute to the size of the observed TC index. For example, if all of the rare
462 species occurred in only the very first or the very last sample, the sample variance in the trend lines
463 would tend to be large compared to that found for the null assemblages. However, less extreme
464 distributions of rare species look very similar to those generated by the null model, and therefore would
465 not contribute substantially to the TC index.

466

467 In the hierarchical model, the assumption of exchangeability of r_i values allowed us to use information
468 from common species to estimate trends for less common species. Nevertheless, when abundance is so
469 low that there are fewer than 10 individuals counted in 14 or more consecutive annual samples,
470 estimating temporal trends with any statistical model is a dubious exercise. For these cases, auxiliary
471 information, stratified sampling, and additional data may be necessary (Dixon *et al.* 2005).

472

In summary, quantifying temporal trends in species abundances is an important forecasting problem. Given the accelerating rates of habitat alteration and global climate change, the strong heterogeneity that we detected in the stream fish and grassland insect data sets (Figures 5 and 7) may be typical; it seems unlikely to us that most long-term temporal trends will be accounted for entirely by the sampling effects in our null model. In these cases, the hierarchical models provide a sensible framework for predicting what the future may hold.

ACKNOWLEDGMENTS

We thank Anne Chao and 2 anonymous reviewers for comments that improved the manuscript. NJG was supported by the U.S. National Sciences Foundation (NSF DEB-0541936) and the U.S. Department of Energy (022821). GDG was supported by the Warnell School of Forestry and Natural Resources and U.S. Department of Agriculture (USDA) Forest Service McIntire-Stennis program grant GEO-00144-MS. Portions of this work represent a contribution from the Harvard Forest Long-Term Ecological Research Site, supported by NSF 06-20443. This work was also facilitated by meetings at the Binary Matrices Working Group at the National Institute for Mathematical and Biological Synthesis, sponsored by NSF, the U.S. Department of Homeland Security, and the USDA through NSF Award #EF-0832858, with additional support from The University of Tennessee, Knoxville.

Appendix A. R-script file for null model analysis.

```
# This program tests for heterogeneity in linear trend lines
of species abundances.

########################################
# GLOBAL VARIABLES

ObservedMatrix <- as.matrix(read.csv("Grossman Streamfish Data.csv", header=TRUE, row.names=1))
# input has species as rows, columns as times, with labels in the first row and column


Time <- c(11, 21, 34, 35, 45, 47, 57, 59, 71, 81, 82, 93, 117, 129, 141, 153)
# these are months since start of 1962 in Grossman streamfish datafile

NRep <- 1000
# this is the number of random replicates to use in the null model analysis

SimVar <- numeric(NRep)
# this is the vector of simulated variances in growth rates

ObsVar <- 0
# this is the single observed variance in growth rates for comparison with null
########################################
# FUNCTION SourcePool
# This function takes an input matrix and a calculated number of missing species
# It returns a probability vector for creating null matrices
# For the missing species we assume their frequency = 0.5*rarest observed species

SourcePool <- function(FMatrix, FMissingSpecies) {

      SourcePool <- rowSums(FMatrix)

 SourcePool <- SourcePool/sum(SourcePool)
 SourcePool <- c(SourcePool,rep(0.5*min(SourcePool),FMissingSpecies))

SourcePool <- SourcePool/sum(SourcePool)
}
########################################
#
#
```

```
#
#
########################################
# FUNCTION RandomSampler
# This function takes as input the SourcePool vector and the original data matrix
# It returns a single null matrix created by sampling individuals one at a time

RandomSampler <- function(FSourcePool, FMatrix) {

TimeProbs <- colSums(FMatrix)/sum(FMatrix)
PoissonN <- numeric(ncol(FMatrix))

for (j in 1:sum(FMatrix)) {
i <-sample(seq(1:length(PoissonN)),1,replace = TRUE,TimeProbs)
PoissonN[i] <- PoissonN[i] + 1
}


SampleMatrix <- matrix(0,length(FSourcePool),ncol(FMatrix))

for (j in 1:ncol(FMatrix)) {
      for (n in 1:PoissonN[j]) {
      i <- sample(seq(1:length(FSourcePool)),1,replace = TRUE,FSourcePool)
SampleMatrix[i,j] <- SampleMatrix[i,j] + 1
}
}
RandomSampler <- SampleMatrix
}
########################################
#
#
#
#
########################################
# FUNCTION GrowthVar
# This function takes as input a data matrix and a vector of times
# It fits a linear model to temporal trends for each species and
# calculates the variance among species in the measured slopes
# The calculation is only made for species that occurred at least once in the matrix
# This function can be modified to fit other kinds of temporal models

GrowthVar <- function(FMatrix, FTime) {

```

```
      Growth <- NULL #creates empty vector of unspecified length

      for (i in 1:nrow(FMatrix)) {

      if (sum(FMatrix[i,]) > 0)
{
      model <- lm(FMatrix[i,]~FTime) #fits model for species that are present
      Growth[i] <- model$coefficients[2] #stores regression slope parameter

}

}


var(Growth, use = "complete") #species with 0 generate missing values that are ignored
}
########################################
#
#
#
#
########################################
# FUNCTION Chao2
# This function takes as input the observed matrix
# It calculated the bias corrected version of Chao2 (Equation 4 in EstimateS manual)
# Note that is this function is needed only once to get missing species for the simulation
# Results are rounded to the nearest whole integer

Chao2 <- function(FMatrix) {
Occurrences <- NULL
for (i in 1:nrow(FMatrix)) {
Occurrences[i] <- sum(FMatrix[i,] > 0)

 }
Uniques <- sum(Occurrences == 1)
Duplicates <- sum(Occurrences == 2)
m <- ncol(FMatrix)   # m = number of samples


Chao2 <- ((m - 1)/m)* ((Uniques*(Uniques - 1)))/(2*(Duplicates + 1))
Chao2 <- round(Chao2)

}
```

```
########################################
#
#
#
#
########################################
# PROGRAM Basic Simulation Loop

MissingSpecies <- Chao2(ObservedMatrix)
SourcePoolVector <- SourcePool(ObservedMatrix,MissingSpecies)
ObsVar <-GrowthVar(ObservedMatrix,Time)

for (i in 1:NRep) {
RandomMatrix <- RandomSampler(SourcePoolVector,ObservedMatrix)
SimVar[i] <-GrowthVar(RandomMatrix,Time)
}
ConfidenceInterval <- quantile(SimVar,c(0.025,0.975))
ConfidenceInterval
ObsVar
SES <- (ObsVar - mean(SimVar))/sqrt(var(SimVar))
SES
sum(ObsVar > SimVar)/NRep
summary(SimVar)
########################################
```

Appendix B. WinBUGS script file for simulation of prior distributions and fit of marginal trend model.

```
model {


#  priors

beta ~ dnorm(0, 0.0001)
sigma ~ dunif(0,10)
tau <- pow(sigma, -2)
for (i in 1:S) {
  a[i] ~ dnorm(0, 0.0001)
}



#  marginal model of trend

for (i in 1:S) {

  r[i] ~ dnorm(beta, tau)

  for (j in 1:T) {

    y[i,j] ~ dpois(mu[i,j])
    log(mu[i,j]) <- a[i] + r[i]*t[j]
  }
}

}
```

669 # LITERATURE CITED

670 Chao, A. 1984 Non-parametric estimation of the number of classes in a population. *Scandinavian*
671 *Journal of Statistics* 11, 265-270.

672 Chao, A., Colwell, R. K., Lin, C.-W. & Gotelli, N.J. 2009 Sufficient sampling for asymptotic minimum
673 species richness estimators. *Ecology* 90, 1125-1133.

674 Clarke, K. R. 1993 Non-parametric multivariate analyses of changes in community structure. Australian
675 *Journal of Ecology* 18, 117-143.

676 Collins S. L., Micheli, F. & Hartt, L. 2000 A method to determine rates and patterns of variability in
677 ecological communities. *Oikos* 91, 285-293.

678 Colwell, R. K. 2009 EstimateS: Statistical estimation of species richness and shared species from
679 samples. Version 8.2. User's Guide and application published at: http://purl.oclc.org/estimates.

680 Colwell, R. K. & Coddington, J. A. 1994 Estimating terrestrial biodiversity through extrapolation. *Phil.*
681 *Trans. Royal Soc. London B* 345, 101-118.

682 Connell, J. H. & Slatyer, R. O. 1977 Mechanisms of succession in natural communities and their role in
683 community stability and organization. *American Naturalist* 111, 1119-1144.

684 Dixon, P. M., Ellison, A. M. & Gotelli, N. J. 2005 Improving the precision of estimates of the
685 frequency of rare events. *Ecology* 86, 1143-1123.

686 Dorazio, R. M., Kery, M., Royle, J.A. & Plattner, M. 2010 Models for inference in dynamic
687 metacommunity systems. *Ecology* (in press).

688 Dunson, W. A. & J. Travis, J. 1991 The role of abiotic factors in community organization. *American*
689 *Naturalist* 138, 1067-1091.

690 Ellison, A. M., Bank, M. S., Clinton, B. D., Colburn, E. A., Elliott, K., Ford, C. R., Foster, D. R.,
691 Kloeppel, B. D., Knoepp, J. D., Lovett, G. M., Mohan, J., Orwig, D. A., Rodenhouse, N. L., Sobczak,
692 W. V. ,Stinson, K. A., Stone, J. K., Swan, C. M., Thompson, J., von Holle, B. & Webster, J. R. 2005
693 Loss of foundation species: consequences for the structure and dynamics of forested ecosystems.
694 *Frontiers in Ecology and the Environment* 3, 479-486.

695 Fujiwara, M. & Mohr, M. S. 2009 Identifying environmental signals from population abundance data
696 using multivariate time-series analysis. *Oikos* 118,1712-1720.

697 Gilks, W. R., Thomas, A. &Spiegelhalter, D. J. 1994 A language and program for complex Bayesian
698 modelling. *The Statistician* 43, 169–178.

699 Gotelli, N.J. & Graves, G.R. 1996 *Null models in ecology*. Smithsonian Institution Press, Washington,
700 DC.

701 Gotelli, N. J. & Colwell, R.K. 2001 Quantifying biodiversity: procedures and pitfalls in the
702 measurement and comparison of species richness. *Ecology Letters* 4, 379-391.

703 Grossman, G. D., Moyle, P. B. & Whittaker, J. O., Jr. 1982 Stochasticity in structural and functional
704 characteristics of an Indiana stream fish assemblage: a test of community theory. *American Naturalist*
705 120, 423-454.

706 Grossman, G. D., Freeman, M. C., Moyle, P. B. & Whittaker, J. O., Jr. 1985 Stochasticity and
707 assemblage organization in an Indiana stream fish assemblage. *American Naturalist* 126: 275-285.

708 Grossman, G. D., Ratajczak, R. E., Crawford, M. K. & Freeman. M. C. 1998 Assemblage organization
709 in stream fishes: effects of environmental variation and interspecific interactions. *Ecological*
710 *Monographs* 68, 395-420.

711 Grossman, G. D. & Sabo, J. L. 2010 Incorporating environmental variation into models of community

712 stability: examples from stream fish assemblages. In *Community Ecology of Stream Fishes* (ed. K. Gido
713 & D. Jackson)Washington DC, USA: American Fisheries Society.

714 Hurlbert, S. H. 1971 The nonconcept of species diversity: a critique and alternative parameters. *Ecology*
715 52, 577-585.

716 Jones, C. G., Lawton, J. H. & Shachak, M. 1994 Organisms as ecosystem engineers. *Oikos* 69, 373-386.

717 Kery, M., Dorazio, R. M., Soldaat, L., van Strien, A., Zuiderwijk, A. & Royle, J. A. 2009 Trend
718 estimation in populations with imperfect detection. *Journal of Applied Ecology* 46, 1163–1172.

719 Kellogg Biological Station Sampling Protcols. 1995 http://lter.kbs.msu.edu/protocols/36

720 Lawton, J. H., Bignell, D.E. & Bolton, B. 1998 Biodiversity inventories, indicator taxa, and effects of
721 habitat modification in tropical rain forest. *Nature* Bignell, D.E. & Bolton, B. (, 72-76.

722 Longino, J. T., Coddington, J. & Colwell, R. K. 2002 The ant fauna of a tropical rain forest: estimating
723 species richness three different ways. *Ecology* 83, 689–702

724 LTER 2007 The US Long Term Ecological Research Network. http://www.lternet.edu/

725 Lunn, D. J., Thomas, A., Best, N. & Spiegelhalter, D. 2000 WinBUGS -- a Bayesian modelling
726 framework: concepts, structure, and extensibility. Statistics and Computing, 10:325-337.
727 http://www.mrc-bsu.cam.ac.uk/bugs/

728 Manchester, S. J. & Bullock, J. M. 2000 The impacts of non-native species on UK biodiversity and the
729 effectiveness of control. *Journal of Applied Ecology* 37, 845-864.

730 Manly, B. J. 2009 *Randomization, bootstrap and Monte Carlo methods in biology*. 3rd edition. Boca
731 Raton, Florida: Chapman and Hall/CRC.

732 McGill, B. J. *et al.* 2007. Species abundance distributions: moving beyond single prediction theories to

733 integration within an ecological framework. *Ecology Letters* 10, 995-1015.

734 Mills, L. S., Soule, M. E. & Doak, D. F. 1993 The keystone species concept in ecology and
735 conservation. *BioScience* 43, 219-224.

736 Pauly, D. 1995 Anecdotes and the shifting baseline syndrome of fisheries. *Trends in Ecology and*
737 *Evolution* 10, 430.

738 R Development Core Team 2008 *R: A language and environment for statistical computing*. Vienna,
739 Austria: R Foundation for Statistical Computing.

740 Robert, C. P. & Casella, G. 2004 *Monte Carlo statistical methods* (second edition). New York:
741 Springer-Verlag.

742 Royle, J. A. & Dorazio, R. M. 2008 *Hierarchical modeling and inference in ecology*. Amsterdam:
743 Academic Press.

744 Sanders, H. 1968. Marine benthic diversity: a comparative study. *American Naturalist* 102, 243-282.

745 Simberloff, D. 1978 Use of rarefaction and related methods in ecology. In*: Biological Data in Water*
746 *Pollution Assessment: Quantitative and Statistical Analyses* (eds K. L. Dickson, J. Cairns Jr. & R.J.
747 Livingston), pp. 150-165. Philadelphia, PA, U.S.A: American Society for Testing and Materials.

748 Solow, A. R. 1994 Detecting change in the composition of a multi-species community. *Biometrics* 50,
749 556-565.

750 Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J. C., Fromentin, J.-M.,
751 Hoegh-Guldberg, O. & Bairlein, F. 2002 Ecological responses to recent climate change. *Nature* 416,
752 389-395.

753 White, E. M., Wilson, J. C. & Clarke, A. R. 2006 Biotic indirect effects: a neglected concept in

754 invasion biology. *Diversity and Distributions* 12, 443-455.

755 Whittaker, J. O. 1976 Fish community changes at one Vigo County locality over a twelve-year period.

756 *Proceedings of the Indiana Academy of Sciences* 85, 191-207.

757

758 Table 1. Empirical details of two data matrices and null model results for temporal change analysis. See
759 text for definitions of variables.

Source	Grossman *et al.* (1982)	KBS (1995)
Taxon	Stream fishes	Grassland insects
Study site	Central Illinois stream reach	Successional grassland plot
Sampling method	Seining	Sticky Traps
Observed number of species	55	9
Estimated number of undetected species (Chao 1984)	16	0
Sampling interval	~ annual	annual
Timespan	September 1963-September 1974	1989-2002
Number of sampling dates (T)	15	14
Average abundance per sample (minimum, maximum)	914 (87, 5344)	401 (71, 1152)
Average abundance per species (minimum, maximum)	266 (1,4304)	624 (2, 2793)
Total abundance (N)	14142	5614
Observed Temporal Change index (TC)	0.256	156.90
Average of 1000 simulated values of TC (95% confidence interval)	0.095 (0.082, 0.111)	39.16 (31.81, 47.22)

P(observed $\overline{TC}$ | null model)	< 0.001	< 0.001

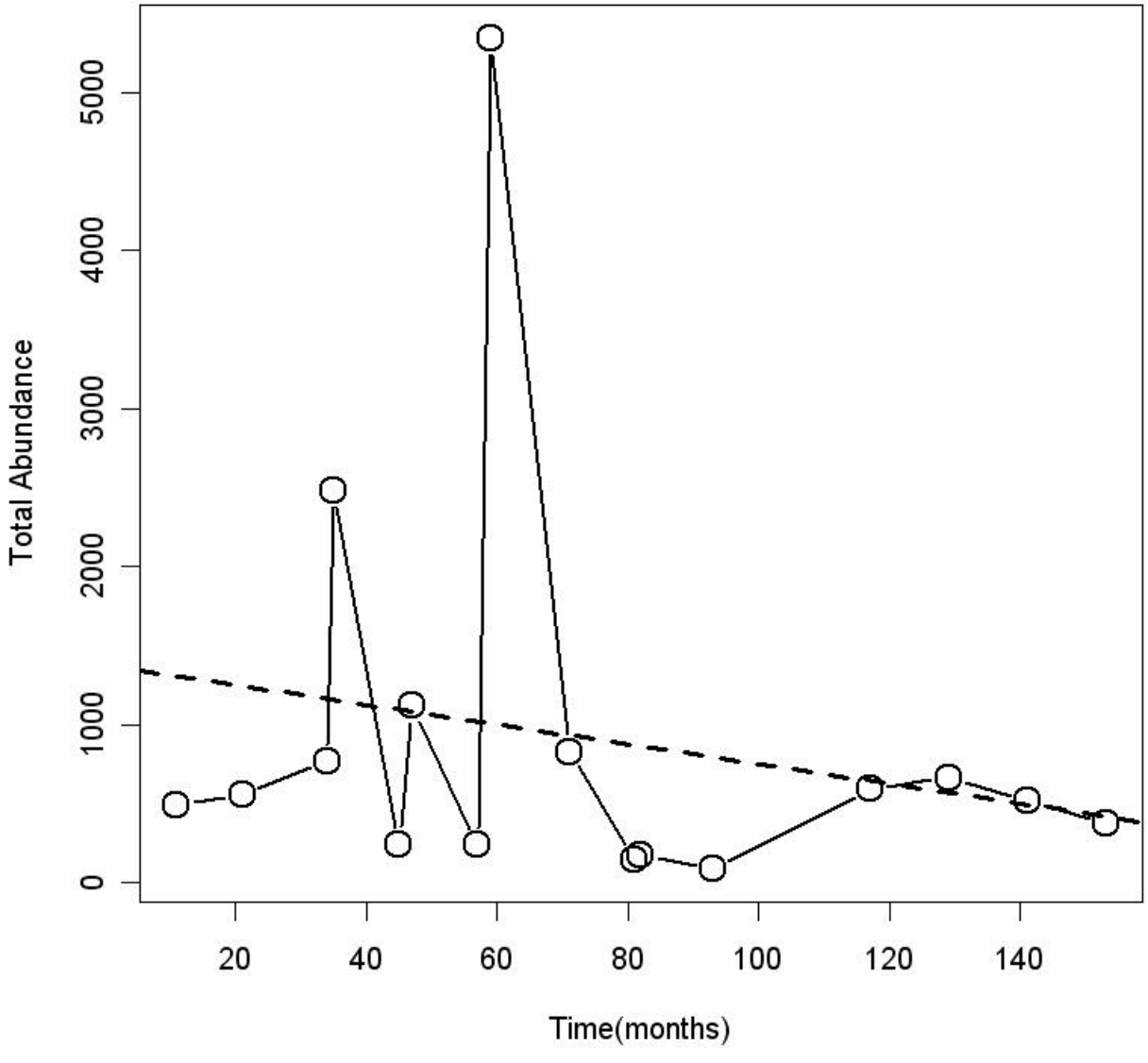
760

761

762 Figure 1. Temporal trends in total abundance for the stream fish samples of Grossman *et al.* (1982). The

763 dashed line indicates the regression line for a simple linear model ($n_t = 1371 - 6.212 t_j$; $r^2 = 0.04$; $P =$

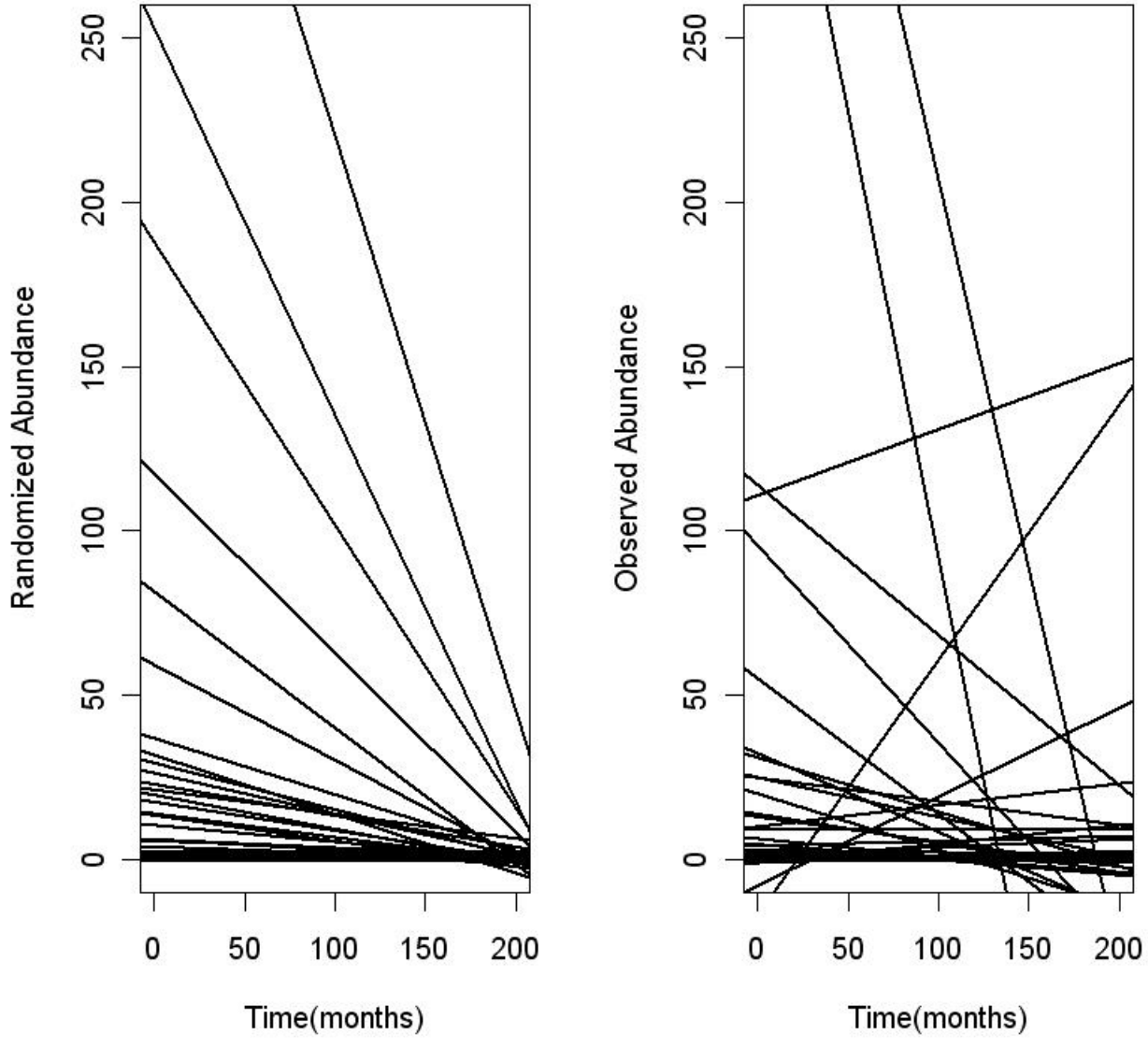
764 0.446)



765

766

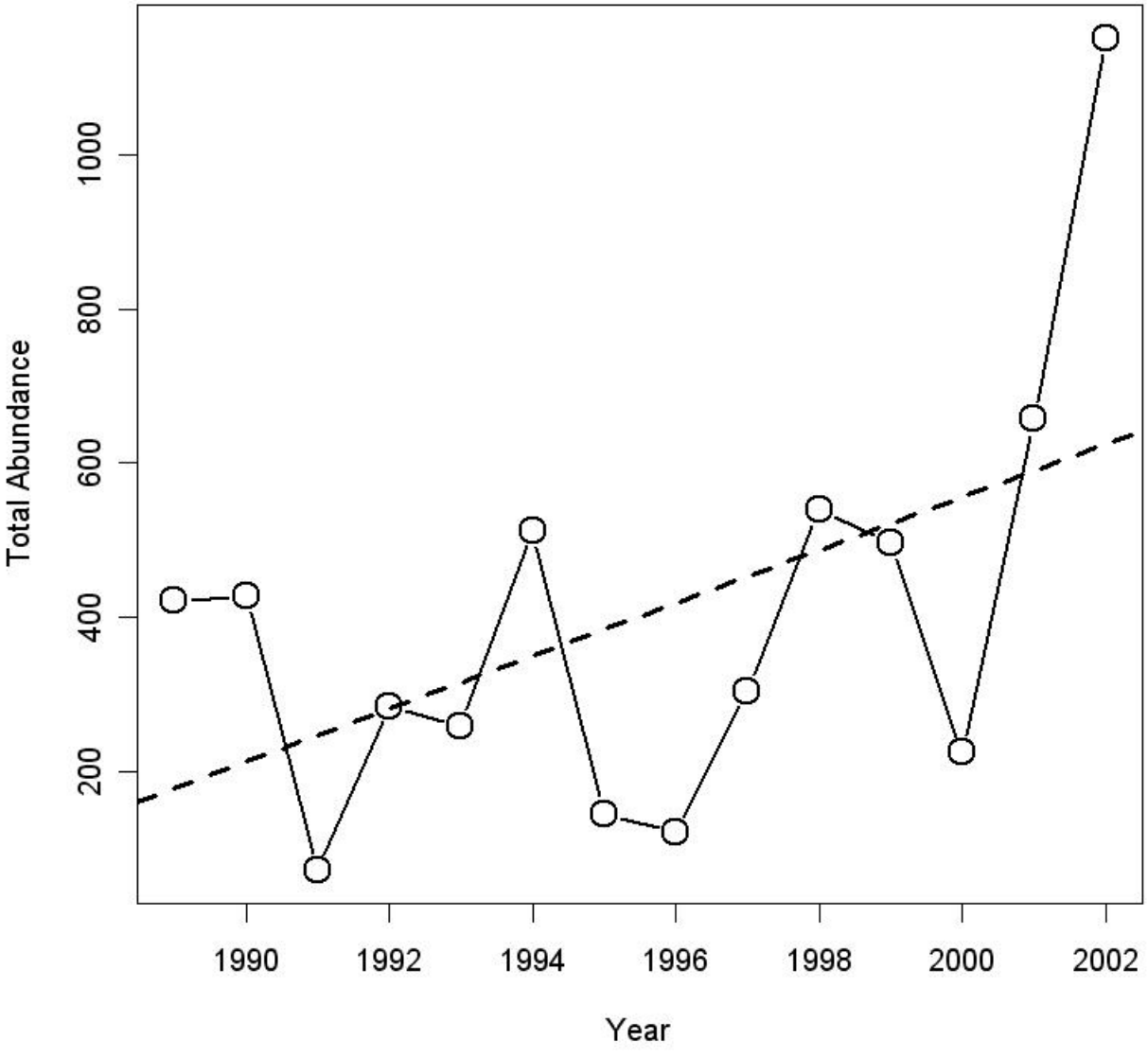
Figure 2. Observed and simulated trend lines for the stream fish data of Grossman *et al.* (1982). Left panel: results of a single replicate of the null model simulation, in which total abundances for each species are sampled randomly from the abundance distribution of all species pooled through time. See text for details of the simulation model. Each line is the least-squares regression for one of the simulated species. Right panel: same graph for the observed data.



775

776 Figure 3. Temporal trends in total abundance for the insect samples of the successional plot. The dashed

777 line indicates the regression line for a simple linear model (n_t= -68253.0 + 34.4t_j; r^2 = 0.27; P = 0.057).



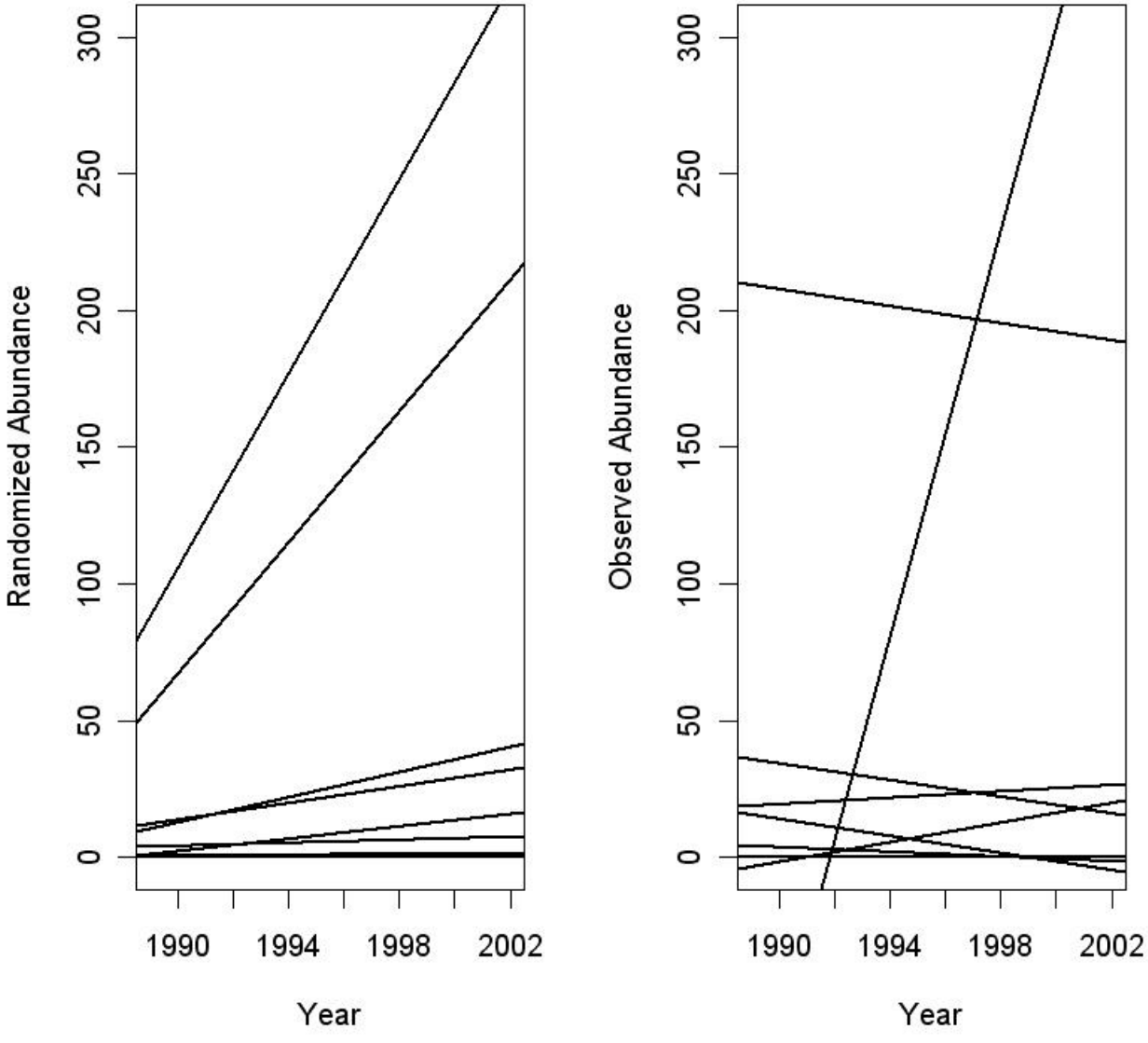
778

779

780

781 Figure 4. Observed and simulated trend lines for the insect data from the successional plot. Details as in

782 Figure 2.

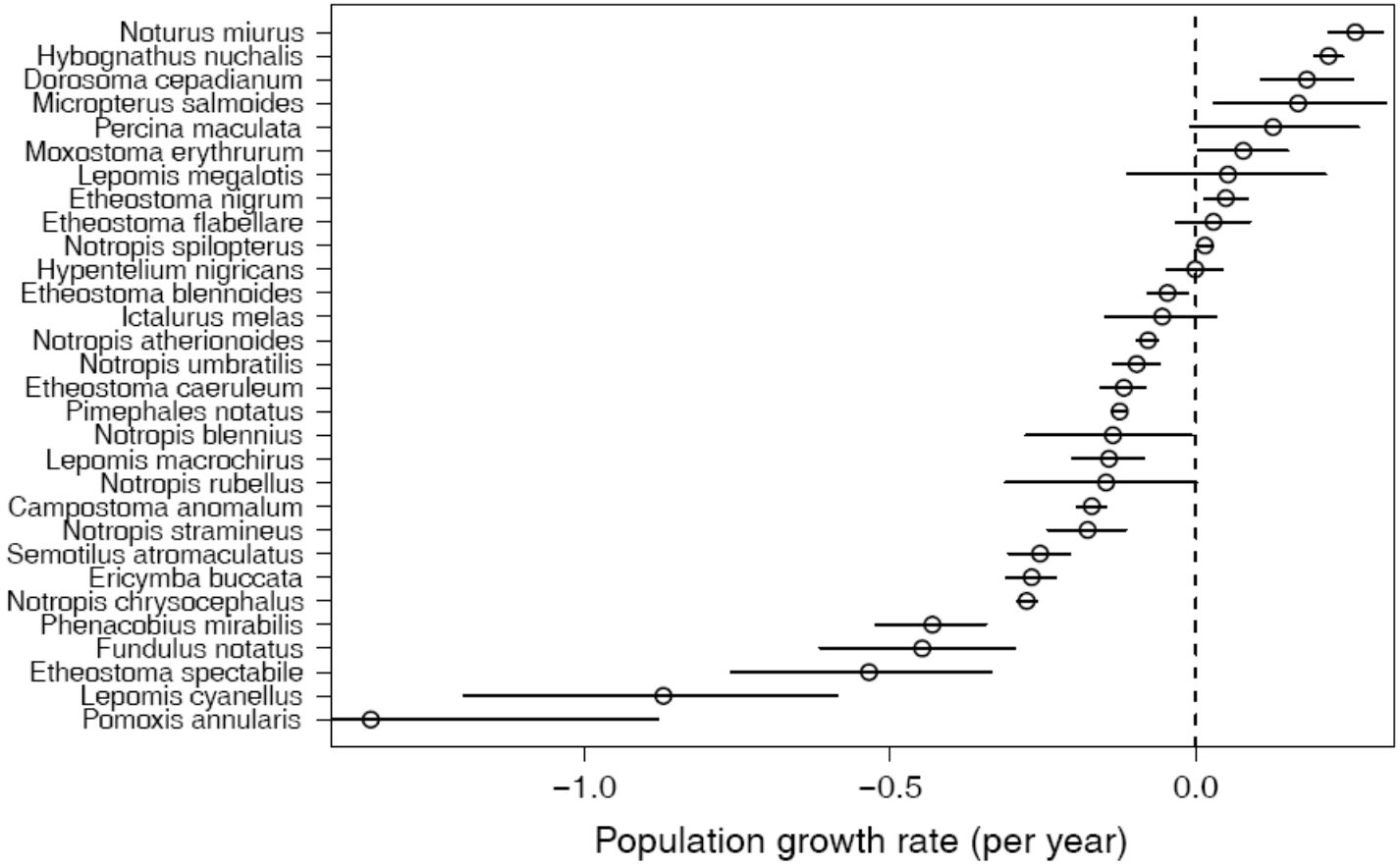


783

784

785 Figure 5. Hierarchical model estimates of r_i, the intrinsic rate of increase (= ln(λ)) for 30 species of stream fishes. Each circle represents the

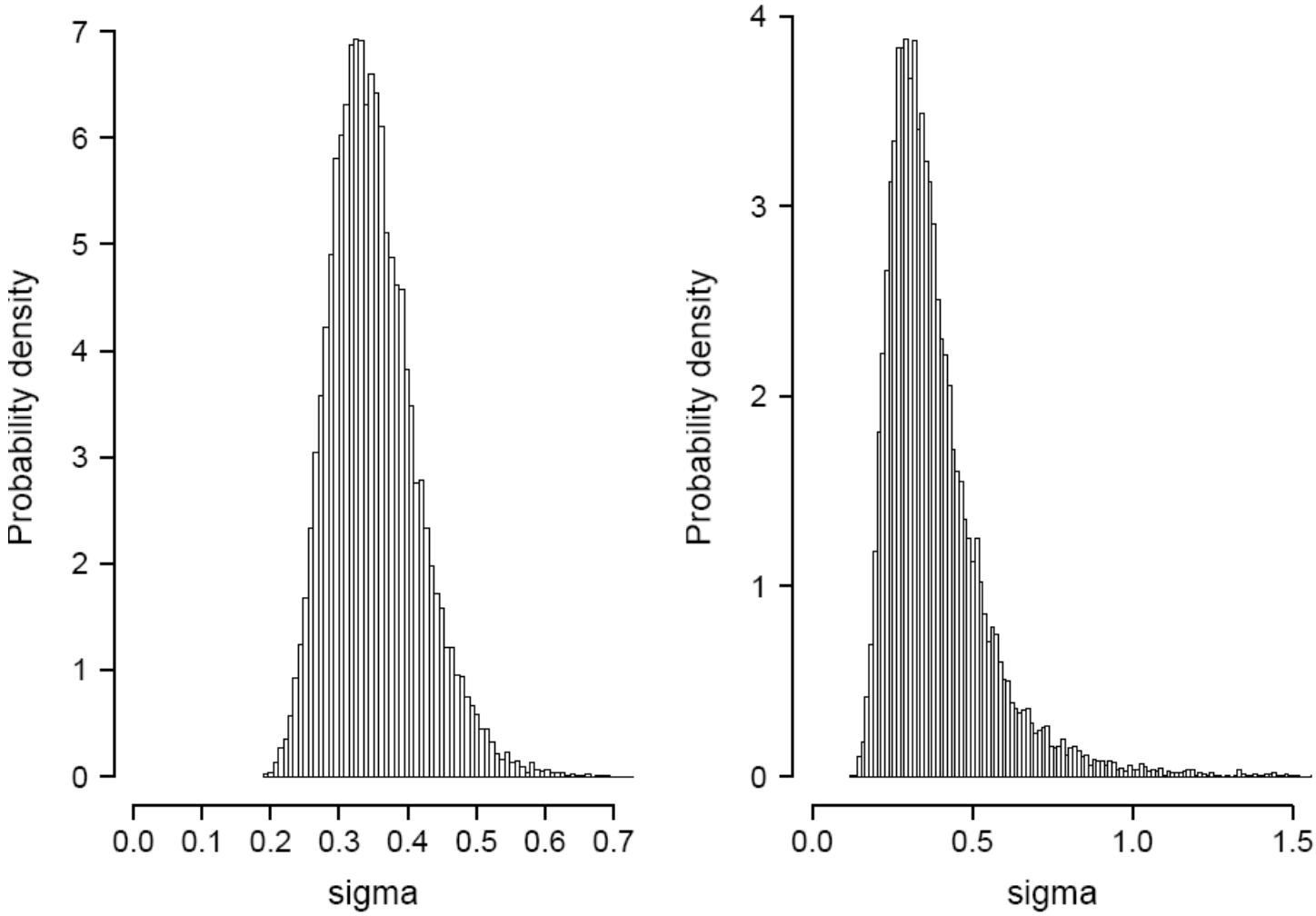
786 estimated r_i, and the error bar is the asymmetrical 95% credible interval.



787

788

Figure 6. Posterior distribution of σ, the variation among species in temporal trends, estimated from the hierarchical model (Eq. 19) by Markov chain Monte Carlo simulation. Left panel: stream fishes. Right panel: grassland insects.


Probability density
0
1
2
3
4
5
6
7
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
sigma
Probability density
0
1
2
3
4
0.0
0.5
1.0
1.5
sigma

792

Figure 7. Hierarchical model estimates of r_i, the intrinsic rate of increase (= ln(λ)) for 7 species of grassland insects. Details as in Figure 5.

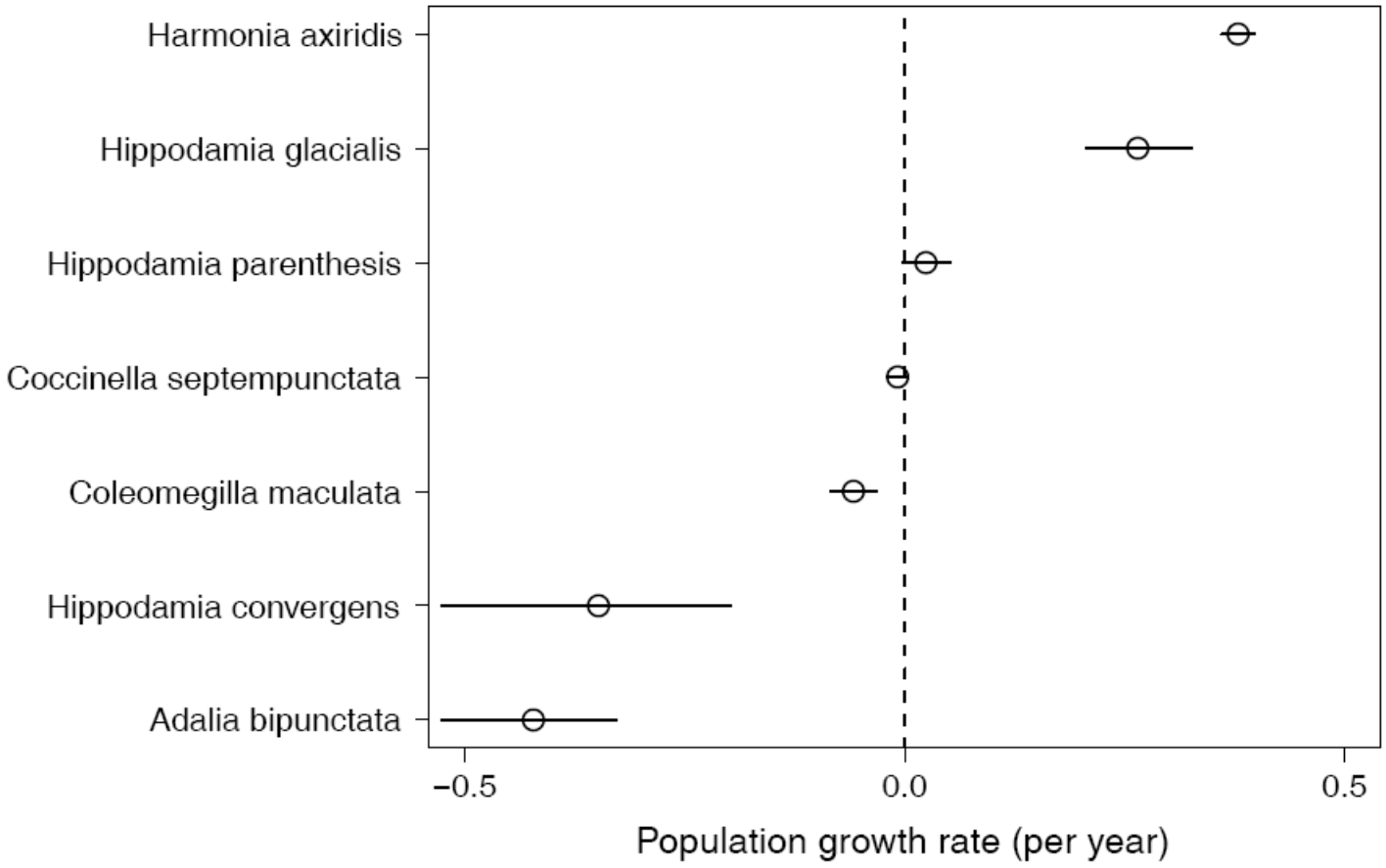


Figure 8. Estimated trends for representative stream fish and insect species. Upper panel: Estimated trends in captures of stream fish (shown by lines) of two species superimposed on the observed counts of these species. Ordinate is logarithmically scaled. Three zero-valued counts (for years 1964, 1966, and 1968) of *Noturus miurus* are not shown. Lower panel: Trends for insect species. Five zero-valued counts (for years 1989-1993) of *Harmonia axiridis* are not shown.